The Effects of Interaction on Bayesian Reasoning

Type: Technique

Saahil Claypool

1 Introduction

Inference has long been a hard problem for humans, regardless of education. But, as evidence based reasoning becomes more important, a flexible system for reasoning about Bayesian statistics is necessary.

While researchers have developed a large number of different techniques to help users understand Bayesian models, the techniques vary in both effectiveness, generalizability, or scalability. Prior effective techniques can be categorized into two (non-distinct) categories: frequency visualizations and structured textual representations [5]. But, frequency diagrams have had contradictory results when generalized to larger populations, and textual representations may be more problem-specific. It may be harder to write a paragraph explaining a problem in a structured manner, whereas graphs can be recreated from the data without very little effort.

In this paper, we discuss the use of interactive visualizations and a question-style presentation of information to help end users make sense of Bayesian models. We postulate that, because only specific portions of the problem space need to be known by the user to answer questions about the Bayesian models [3], we can reduce the cognitive load on the user by constraining the information presented to the user with interaction and probing questions. While prior work has been in interactive Bayesian representations, those visualizations do not constrain the problem, and still force the user to keep a larger portion of the Bayesian model in memory. Our visualizations aim to fix this by showing only pertinent information to the users' queries.

2 One-sentence description

To create an easier to understand representation for Bayesian probability models, we explore the user of interaction to reduce the information presented to a user, and thus improve their ability to make accurate inferences.

3 Project Type

Technique

The goal would be to create a new technique and evaluate how effective it is for Bayesian Inference tasks.

4 Audience

The outcomes of this project will help researchers understand how to better present Bayesian probabilities, and also understand how the selective presentation of information can help improve task based activities.

Without this work, there will continue to be a large number of false assumptions caused by a misunderstanding of Bayesian probabilities.

5 Approach

5.1 Details

1. Create a small interactive tool

   The 'tool' will be provided as a reference similar to prior studies. This will hopefully help the user reason about the problem by providing a visual aid. So, while the numbers will be different, it should allow them to make a similar 'query' and thus answer their question about their sample.

2. Quiz the user

   Given the interactive template or tool, the user will be asked some Bayesian question about a new sample. This will be compared to a set of users without the interactive tool.

The question will asked similar to prior work:

> If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also get a positive mammography. A woman in this age group had a positive mammography in a routine screening [3].

5.2 Evidence for Success

indicates that people do not need to keep the entire Bayesian model in their head to make inferences, but rather just the few pertinent numbers for their estimations [3]. So, by allowing people to narrow down the Bayesian model to answer just the question asked, maybe they will be able to ignore the other distracting numbers and better answer the question.

This interactive design has been tested before [6], but failed to reach significant results. This could be because the visualizations provided did not hide extraneous information, nor did it use aligned axes for the comparison of categories. Both of these together may have made it harder for users to reason about the problem.

Further, with modern technologies like JavaScript, we should be able to create a richer interactive model that is more natural for the user to use than an excel sheet and visual basic scripts.

6 Best-case Impact Statement

This paper will be successful if it is able to provide definitive results on an interactive design helping in Bayesian reasoning.

7 Major Milestones

1. Create the interactive tool
2. Create a testing framework to run users through

   including the data and questions they will be asked about the data
3. Run a small set of pilot users

   Tweak study as necessary
4. Run a larger set of users
5. OR run a case study detailing how a smaller set of users worked with the interactive representation

   This will be a backup study if we cannot feasibly run a user study with enough participants.

8 Obstacles

8.1 Major obstacles

1. Cannot get users within the next 7 weeks
2. Users cannot figure out how to interact with a tool on a website

   There isn't a guarantee that non-technical users will intuitively click around on a graph.

8.2 Minor obstacles

1. Scaffolding the experiment is slightly undefined

9 Resources Needed

1. A lab to run users
2. OR Mechanical Turk

10 5 Related Publications

1. How to improve Bayesian reasoning [3]
2. Assessing the effect of visualizations on Bayesian reasoning through crowdsourcing [4]
3. Improving Bayesian Reasoning: The Effects of Phrasing, Visualization, and Spatial Ability [5]
4. Interactive visualizations to improve Bayesian reasoning [6]
5. Pictorial representations in statistical reasoning [1]
6. Using tree diagrams without numerical values in addition to relative numbers improves students' numeracy skills: A randomized study in medical education [2]

11 Define Success

Optimally, we will have (statistically) significant results indicating groups with an interactive model will out-perform those with just text.

We may be able to make do with case studies of smaller users and a discussion of how they ended up actually interacting (or not) with the interactive model. This could provide insight on how untrained

users are to interact with a computer to solve problems. These case studies would provide a lower barrier for a useful outcome of our paper.

References

[1] G. L. Brase. Pictorial representations in statistical reasoning. *Applied Cognitive Psychology*, 2009.

[2] H. Friederichs, S. Ligges, and A. Weissenstein. Using tree diagrams without numerical values in addition to relative numbers improves students' numeracy skills: A randomized study in medical education. *Medical Decision Making*, 2014.

[3] G. Gigerenzer and U. Hoffrage. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 1995.

[4] L. Micallef, P. Dragicevic, and J. D. Fekete. Assessing the effect of visualizations on bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics*, 2012.

[5] A. Ottley, E. M. Peck, L. T. Harrison, D. Afergan, C. Ziemkiewicz, H. A. Taylor, P. K. Han, and R. Chang. Improving Bayesian Reasoning: The Effects of Phrasing, Visualization, and Spatial Ability. *IEEE Transactions on Visualization and Computer Graphics*, 2016.

[6] J. Tsai, S. Miller, and A. Kirlik. Interactive visualizations to improve Bayesian reasoning. In *Proceedings of the Human Factors and Ergonomics Society*, 2011.